

# Multi-Dimensional Inheritance

Gregor Erbach

\*Universität des Saarlandes  
FR 8.7 Computerlinguistik  
D-66041 Saarbrücken, Germany  
erbach@coli.uni-sb.de

(cmp-lg/9411025)

April 1994

## Abstract

In this paper, we present an alternative approach to multiple inheritance for typed feature structures. In our approach, a feature structure can be associated with several types coming from different hierarchies (dimensions). In case of multiple inheritance, a type has supertypes from different hierarchies. We contrast this approach with approaches based on a single type hierarchy where a feature structure has only one unique most general type, and multiple inheritance involves computation of greatest lower bounds in the hierarchy. The proposed approach supports current linguistic analyses in constraint-based formalisms like HPSG, inheritance in the lexicon, and knowledge representation for NLP systems. Finally, we show that multi-dimensional inheritance hierarchies can be compiled into a Prolog term representation, which allows to compute the conjunction of two types efficiently by Prolog term unification.

## 1 Introduction

In the past years, multiple inheritance has been increasingly used for the description of natural languages. Some examples are the work in Head-driven Phrase Structure Grammar (HPSG) on the structure of the lexicon [13, 19], semantic sorts for selectional restrictions [5], derivational morphology [20, 15], and the syntax of English relative clauses [21]. As an example of this type of analysis, figure 1 shows a recent HPSG description of English clauses.<sup>1</sup>

In the example, relative clauses are cross-classified according to two “dimensions”; on the one hand according to the Phrase-Type (Headed or Non-Headed), and on the other hand according to the Clause-Type (Interrogative, Declarative, or

---

<sup>\*</sup>This work was supported by the Deutsche Forschungsgemeinschaft, Special Research Division 314 “Artificial Intelligence - Knowledge-Based Systems” through project N3 “Bidirectional Linguistic Deduction” (BiLD), the Commission of the European Communities through the project LRE-61-061 “Reusable Grammatical Resources,” and Cray Systems. Part of the work was done during a visit at the Human Communication Research Centre, University of Edinburgh. I would like to thank the members of the Language Technology Group at HCRC, Chris Brew, Suresh Manandhar, Drew Moshier and Hans Uszkoreit for their comments.

<sup>1</sup>This figure is taken from a slide presented by Ivan Sag at the 1993 EACL meeting in Utrecht. The use of horizontal lines in the hierarchy to indicate a split into several dimensions has been adopted from the book on the Core Language Engine [5].

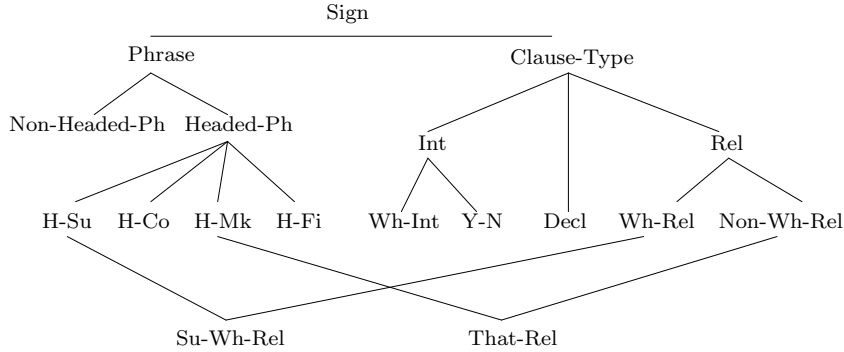


Figure 1: HPSG type hierarchy *with* multi-dimensional inheritance

Relative). The choices within one dimension are mutually exclusive: no structure can be described as both Headed and Non-Headed, or as more than one of {Int, Decl, Rel}. However, a structure can be assigned types from different dimensions, without the need for a subtype that inherits from both dimensions.

In section 2, we present a concept of typing, which provides direct support for multi-dimensional inheritance, and compare it to the type hierarchies in Carpenter’s typed feature logic. Section 3 shows that unification in multi-dimensional inheritance hierarchies can be implemented efficiently as unification of a Prolog term representation of the types. Section 4 applies multi-dimensional inheritance to the problem of systemic classification.

## 2 Multi-Dimensional Inheritance

We follow the Carpenter’s formalisation of typed feature logic, but modify the conception of the type hierarchy. In Carpenter’s logic [8], the type hierarchy is required to be a bounded complete partial order (BCPO), which means that any two types which do have a common subtype must have a unique most general common subtype. A type hierarchy as in figure 1 fails this requirement because the types *Headed-Ph* and *Rel* have two common subtypes, but none of them is more general than the other. In order to make the type hierarchy a BCPO, additional types must be introduced, resulting in a hierarchy like the one in figure 2. In Carpenter’s system, every feature structure has only one unique most general type, so that it is not possible to assign a feature structure two types neither of which subsumes the other unless they have a common subtype.<sup>2</sup>

Carpenter [8, pages17-32] describes a *conjunctive type construction* by which a type hierarchy like the one in figure 1 can be converted into a bounded complete partial order like in figure 2. Once this is done, efficient algorithms for the calculation of greatest lower bounds can be used [2]. We argue that such a conjunctive type construction is neither necessary for theoretical reasons nor for reasons of efficient implementation.

In our system, a feature structure can have different types as long as they are chosen from different dimensions. Our syntax for subtype declarations, given in (1),

<sup>2</sup>A similar point holds for the type system of the logic programming language LIFE [3, 1], with the difference that the unification of two types need not have a unique result; the unification of the types *Headed-Ph* and *Rel* would create a choice point and produce the two alternative solutions *Su-Wh-Rel* and *That-Rel*.

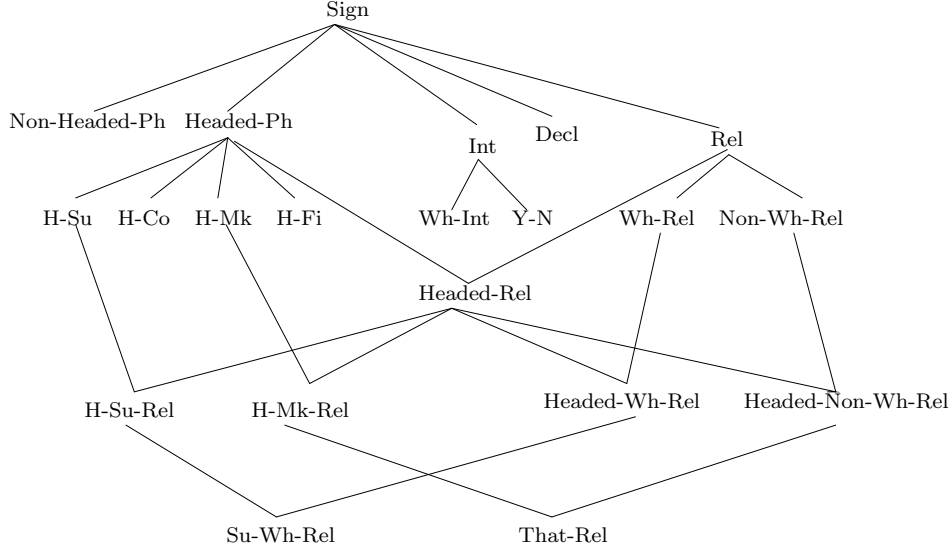


Figure 2: HPSG type hierarchy with one-dimensional inheritance

combines information about subtyping and disjointness. All the  $Y_i$  are subtypes of  $X$ , and all  $Y_i$  are disjoint.

$$X > [Y_1, \dots, Y_n]. \quad (1)$$

Multi-dimensional inheritance arises in the case where there is more than one declaration with the same supertype on the left-hand side, as in (2).

$$\begin{array}{rcl} X & > & [Y_{1.1}, \dots, Y_{1.n}]. \\ & \vdots & \\ & \vdots & \\ X & > & [Y_{m.1}, \dots, Y_{m.k}]. \end{array} \quad (2)$$

Instead of writing a separate subtype declaration for each dimension, multiple dimensions are conventionally connected with the product operator  $*$ , as in declaration (3), which is equivalent to the declarations in (2).

$$X > [Y_{1.1}, \dots, Y_{1.n}] * \dots * [Y_{m.1}, \dots, Y_{m.k}]. \quad (3)$$

Multiple inheritance is the case where some  $Y$  type occurs in the right-hand side of more than one type declaration. In this case, the type has several supertypes, which must be chosen from different dimensions in order to be consistent with each other.<sup>3</sup>

We now turn to the semantics of type declarations. The denotation of each type is a subset of the domain. The semantics of the type declaration in (1) is given by the axioms (4), which states that the denotation of any of the  $Y_i$  is a subset of the denotation of  $X$ , and (5), which states that all the  $Y_i$  are disjoint. No additional axioms are needed in case of multi-dimensional inheritance.

<sup>3</sup>We ignore the case where one of the supertypes subsumes the other since such declarations are redundant.

$$\forall Y_i (\llbracket Y_i \rrbracket \subseteq \llbracket X \rrbracket) \quad (4)$$

$$\forall Y_i \forall Y_j (Y_i \neq Y_j \Rightarrow \llbracket Y_i \rrbracket \cap \llbracket Y_j \rrbracket = \emptyset) \quad (5)$$

Our system has an open-world semantics for type hierarchies. A feature structure can be described by two types from different hierarchies, but there need not be a common subtype of these two types. This is in contrast with a system like ALE [9] with a closed-world semantics where the conjunction of two types is inconsistent unless one subsumes the other or they have a common subtype.

```
sign > [non_headed_ph,headed_ph] * [int,decl,rel].
headed_ph > [h_su,h_co,h_mk,h_fi].
int > [wh_int,y_n].
rel > [wh_rel,non_wh_rel].
h_su > [su_wh_rel].
wh_rel > [su_wh_rel].
h_mk > [that_rel].
non_wh_rel > [that_rel].
```

Figure 3: Type declarations

Our notion of feature typing and appropriateness is based on Carpenter’s feature logic. Every feature is introduced for a unique most general type, and is appropriate for all subtypes of that type. In case of multiple inheritance, a type can inherit different features from its supertypes in different dimensions. A difference arises with type restrictions for feature values. In Carpenter’s system, the value of a feature has one type as the type restriction, whereas in our system, the type restriction can be a conjunction of types from different dimensions.

In our system, the type hierarchy from figure 1 can be expressed directly with the declarations given in figure 3.

Our notion of typing is similar to the one adopted in the *Comprehensive Unification Formalism* (CUF [11]), in that it adopts an open-world semantics, and two types are considered as consistent unless they are explicitly declared to be disjoint. However, CUF allows to state more general type axioms using the full power of propositional logic. In our system, the type axioms are restricted to subtyping (which corresponds to implication) and disjointness. This restriction allows the efficient compilation of multi-dimensional inheritance hierarchies to Prolog terms, which will be described in the next section.

### 3 Compilation into a Prolog Term Representation

Multi-dimensional type hierarchies have the favourable property that the types can be compiled to a Prolog term representation. With this representation, Prolog’s built-in term unification is all that is required to compute the conjunction of two types. The Prolog term representation given here builds on and extends the representation introduced by Mellish [16, 17] and used in the Core Language Engine [5] and the ALEP grammar development system [4, 6].

We start out by describing how the translation of type hierarchies into Prolog terms works, and then give an example. The translation to terms must be able to handle different dimensions of typing, mutually exclusive choices in a dimension, subtyping, multiple inheritance, features, and equality.

**Different dimensions:** Each dimension occupies a different argument position in the resulting term representation, so that information from different dimensions can be combined by unification.

**Mutually exclusive types:** Mutually exclusive types in the same dimension have different functors at the same argument position, so that their unification fails.

**Subtype:** The term which corresponds to the subtype is a further instantiation of the term corresponding to its supertype.

**Multiple inheritance:** The term which corresponds to the subtype is a further instantiation of the unification of the terms which correspond to its supertypes.

**Feature:** The term representation has an argument position for each feature introduced for a type. If a feature is introduced for a subtype, then an argument position is provided in that argument which further instantiates the supertype.<sup>4</sup>

**Equality:** In order to be able to distinguish structures that are identical from those which just happen to have the same value (i.e. their term representation is instantiated to the same ground term), an extra variable is introduced in the term representation (preventing instantiation to a *ground* term), which is only equal for two structures if they have been made identical by unification.

Type	Encoding
Sign	sign(.,.)
Non-Headed-Ph	sign(non_headed,.)
Headed-Ph	sign(headed_ph(,),.)
Int	sign(.,int(.,))
Decl	sign(.,decl)
Rel	sign(.,rel(.,))
H-Su	sign(headed_ph(h_su(,),.),.)
Wh-Rel	sign(.,rel(wh_rel(.,)))
Su-Wh-Rel	sign(headed_ph(h_su(su_wh_rel)),rel(wh_rel(su_wh_rel)))
...	

Figure 4: Encoding of Types as Prolog Terms

Given this kind of representation, two typed feature structures can be unified simply by unification of the corresponding Prolog terms. In figure 4, we provide an encoding of the types in the hierarchy from figure 1 by Prolog terms. For simplicity, we leave out the argument position used to establish the equality of feature structures, and any argument positions used to encode features. Note that the type Su-Wh-Rel makes a choice in both dimensions.

If the type Su-Wh-Rel has any subtypes, a choice must be made which of the two occurrences of `su_wh_rel` in the Prolog terms should get argument positions for carrying this information.<sup>5</sup> We always choose the leftmost occurrence in a term for representing subtypes (and features). Further occurrences then only serve to make a choice in a particular dimension of the hierarchy, and for that purpose, an

<sup>4</sup>The compilation of feature structures to Prolog terms has been described in [22, 14, 10], but these works assume untyped feature structures. One may wonder what the difference is between features and multiple typing dimensions, since they have very similar term representations. A technical answer would be that equality constraints (coreferences) can be stated over feature values, but not over type dimensions. Work by Moshier [18] on a rational reconstruction of typed feature structures in domain theory approaches the question from a more fundamental perspective.

<sup>5</sup>From the correctness of the translation, it would be no problem representing the subtypes in both occurrences, but such an encoding is clearly redundant.

atom which is distinct from other terms that can occur as alternatives in the same dimension is sufficient.

## 4 Application to Systemic Classification

In this section, we apply the conception of multi-dimensional inheritance to systemic classification networks, which have been discussed in [16].

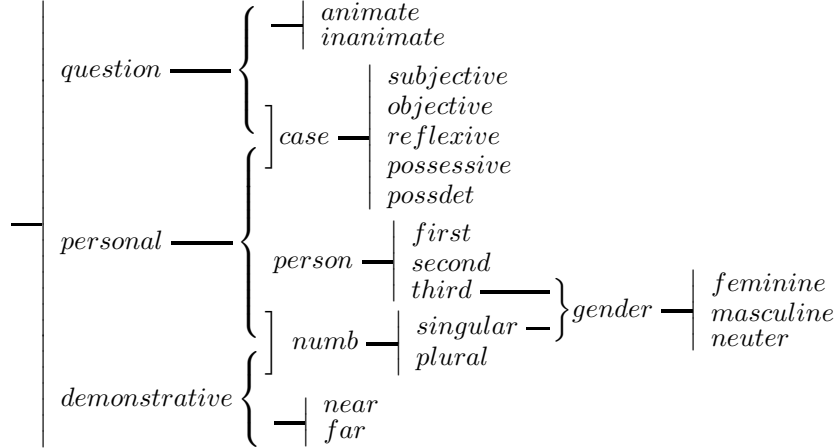


Figure 5: Winograd’s systemic pronoun network

Systemic classification networks are an interesting formalism for such an encoding because they offer considerable expressive power. Figure 5 shows a systemic classification network for English pronouns taken from [23].

Figure 6 shows the connectives, and the translation of the first three in our system. The final connective (disjunctive entry condition) has no simple translation. This is not surprising, given the complexity analysis by Brew [7] who shows systemic classification to be NP-hard by giving an encoding of the 3SAT-problem in a systemic network with disjunctive entry conditions.

We treat disjunctive entry conditions by introduction of new types into the hierarchy. For each type  $X$  at the right-hand side of a disjunctive entry condition, we introduce two new types,  $X'$  for the original type, and  $\neg X$  for its negation. These pairs of new types are introduced in different dimensions at the top of the choice system containing the disjunctive entry conditions. All types  $Y$  which have subtypes that are on the left-hand side of the disjunctive entry condition for  $X$  become subtypes of  $X'$ , and all other types become subtypes of  $\neg X$ . In effect, this is an expansion of the disjunctive entry conditions to disjunctive normal form.

In the worst case, this method can lead to translations which are exponentially larger than the original classification network, as for the 3SAT problem. We show by an example that this need not be the case in practice by converting the classification of pronouns given in figure 5 into our system (figure 7).

In this case, our translation is more efficient than the “brute-force translation” described in [16] for networks with disjunctive entry conditions. In Winograd’s systemic network, there are 54 possibilities (6 for interrogative pronouns, 36 for personal pronouns, and 4 for demonstrative pronouns), resulting in a brute-force translation with 55 arguments.<sup>6</sup>

<sup>6</sup>The “brute-force translation” works by encoding  $n$  possibilities into a term with  $n + 1$  argu-

Connective		Type Declarataions
choice system	$X \left  \begin{array}{c} Y_1 \\ \vdots \\ Y_n \end{array} \right.$	$X > [Y_1, \dots, Y_n]$
multiple choices	$X \left\{ \begin{array}{c} Y_1 \\ \vdots \\ Y_n \end{array} \right.$	$X > Y_1 * \dots * Y_n$
conjunctive entry cond.	$\left. \begin{array}{c} X_1 \\ \vdots \\ X_n \end{array} \right\} Y$	$X_1 > [Y, \dots]$ $\vdots$ $X_n > [Y, \dots]$
disjunctive entry cond.	$\left. \begin{array}{c} X_1 \\ \vdots \\ X_n \end{array} \right] Y$	$Y'$ and $\neg Y$ move to the top of the hierarchy

Figure 6: Systemic network connectives and multi-dimensional inheritance

In contrast, in our type system, the Prolog term translation has at most 8 nodes in the worst case, as in the following term which encodes the subjective case masculine (third person singular) personal pronoun.

```
pronoun(case(personal(third(masc)),subjective),num(personal,sg))
```

## 5 Conclusion

We have presented a concept of inheritance which provides direct support for current linguistic descriptions making use of “cross-classification”, and can be compiled into an efficient Prolog term representation.

Given the need for multiple dimensions in linguistic descriptions, we believe that multi-dimensional type hierarchies will remain important even when their compilation into Prolog terms is not needed any longer because unification of typed feature terms will be built-in in future logic programming languages.

For the time being, however, the combination of multi-dimensional inheritance and compilation into Prolog terms appears to give both the efficiency and the expressive power needed to develop larger-scale grammars and lexicons, and use existing Prolog-based technology (DCG parsers, left-corner, head-corner, or chart parsers, semantic-head driven or tabular generators) to build NLP systems. Such an approach can benefit from all the advantages of modern Prolog compilers (indexing, coroutining facilities, module systems etc.) that would need considerable effort to duplicate in a dedicated grammar formalism.

The multi-dimensional inheritance described in this paper is implemented in the system ProFIT [12], which translates programs containing typed feature terms to

ments, whose first and last argument are different (either instantiated to different atoms or related by an inequality constraint). For encoding a value that excludes the  $n^{th}$  possibility, the  $n^{th}$  and  $n + 1^{st}$  argument position are unified. If a combination of values excludes all possibilities, then all arguments are unified with each other, including the first and the last, which are different, so that the unification fails. There is limited use for this type of translation to avoid creation of choice points in disjunctions with a small number of possible values. This can be regarded as a Prolog implementation of finite domains.

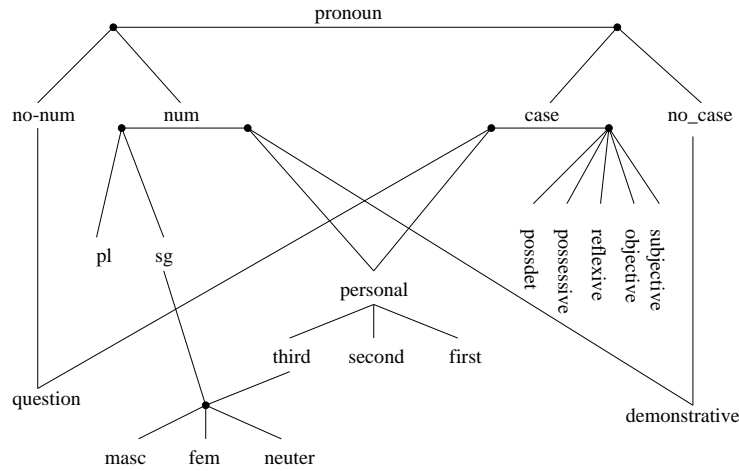


Figure 7: Pronoun hierarchy with disjunctions moved to the top

ordinary Prolog programs.

## References

- [1] Hassan Aït-Kaci and Patrick Lincoln. LIFE, a natural language for natural language. *T. A. Informations*, 30(1-2):37 – 67, 1989.
- [2] Hassan Aït-Kaci and Roger Nasr. LOGIN: A logical programming language with built-in inheritance. In *Proceedings of the 13th ACM Symposium on Principles of Programming Languages*, pages 219–228, St. Petersburg, Florida, 1986.
- [3] Hassan Aït-Kaci and Andreas Podelski. Towards a meaning of LIFE. Technical Report 11, DEC Paris Research Laboratory, Paris, June 1991.
- [4] H. Alshawi, D. J. Arnold, R. Backofen, D. M. Carter, J. Lindop, K. Netter, J. Tsujii, and H. Uszkoreit. Eurotra 6/1: Rule formalism and virtual machine design study. final report. Technical report, SRI International, Cambridge, 1991.
- [5] Hiyan Alshawi, editor. *The Core Language Engine*. MIT Press, 1991.
- [6] BIM-SEMA. *ALEP System Documentation: The ALEP Linguistic Subsystem, Version 1.0*. Commission of the European Communities, March 1993.
- [7] Chris Brew. Systemic classification and its efficiency. *Computational Linguistics*, 17(4):375 – 408, 1991.
- [8] Bob Carpenter. *The logic of typed feature structures*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, Cambridge, 1992.
- [9] Bob Carpenter. *ALE Version β: User Manual*. University of Pittsburgh, 1993.
- [10] Michael Covington. GULP 2.0: an extension of Prolog for unification-based grammar. Technical Report AI-1989-01, Advanced Computational Methods Center, University of Georgia, 1989.



- [11] Jochen Dörre and Michael Dorna. CUF – A formalism for linguistic knowledge representation. In Jochen Dörre, editor, *Computational Aspects of Constraint-Based Linguistic Description. Deliverable R1.2.A*. DYANA-2 – ESPRIT Basic Research Project 6852, 1993.
- [12] Gregor Erbach. *ProFIT User’s Manual - Version 1.05*. Universität des Saarlandes, Saarbrücken, July 1994.
- [13] Dan Flickinger and John Nerbonne. Inheritance and complementation: a case study of easy adjectives and related nouns. *Computational Linguistics*, 18(3):269 – 309, 1992.
- [14] Susan Beth Hirsh. P-PATR: A compiler for unification-based grammars. Master’s thesis, Stanford University, Stanford, CA, December 1986.
- [15] Hans-Ulrich Krieger and John Nerbonne. Feature-based inheritance networks for computational lexicons. In Ted Briscoe, Anne Copestake, and Valeria de Paiva, editors, *Default Inheritance within Unification-Based Approaches to the Lexicon*. Cambridge University Press, 1992.
- [16] Christopher S. Mellish. Implementing systemic classification by unification. *Computational Linguistics*, 14(1):40–51, 1988.
- [17] Christopher S. Mellish. Term-encodable description spaces. In D. R. Brough, editor, *Logic Programming: New Frontiers*, pages 189 – 207. Intellect, Oxford, 1992.
- [18] M. Andrew Moshier. A rational reconstruction of the domain of feature structures. Universität des Saarlandes, Saarbrücken, 1994.
- [19] Carl J. Pollard and Ivan A. Sag. *Information-Based Syntax and Semantics. Volume 1: Fundamentals*. CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA, 1987.
- [20] Susanne Riehemann. Word formation in lexical type hierarchies: A case study of *bar*-adjectives in german. Master’s thesis, University of Tübingen, Department of Linguistics, 1992.
- [21] Ivan Sag. Relative clauses: A multiple inheritance analysis. Stanford, 1994.
- [22] Andreas P. Schöter. Compiling feature structures into terms: A case study in Prolog. Technical Report RP-55, University of Edinburgh, Centre for Cognitive Science, 1993.
- [23] Terry Winograd. *Language as a Cognitive Process*, volume I: Syntax. Addison Wesley, Reading, MA, 1983.